Parallel Task Routing for Crowdsourcing

Jonathan Bragg
UW

Andrey Kolobov
MSR

Mausam
IIT Delhi

Daniel S. Weld
UW

HCOMP  November 3, 2014
Task routing

Workers

Questions
Task routing

Workers

Questions

Skill $\gamma$

Difficultly
Prior work

• Offline
• Sequential routing
• Workers are willing to wait

Parallel routing

1. A pool of workers arrives
2. Assign a task to each worker
3. Workers take the same time to answer

Will relax this assumption
Applications

• Volunteer (e.g., Zooniverse)
• Pay by task (e.g., Amazon Mechanical Turk)
• Pay by time (e.g., oDesk)
Problem space
Problem space

Skill, Difficulty

Latent

Known

Concurrency

Sequential

Parallel
Problem space

- Skill, Difficulty
- Latent
- Known
- Sequential
- Parallel
- Offline
- Adaptive
- Policy
- Concurrency
Problem space

Community-based aggregation
[Venanzi et al. 2014]

Skill, Difficulty

Question features

Policy

Concurrency

Latent

Known

Offline

Sequential

Adaptive

Parallel
Overview

• Workers not willing to wait
• Offline algorithm Juggler_{OFF}
• Adaptive algorithm Juggler_{AD}
  – Scalability
  – Variable response times
Problem definition

<table>
<thead>
<tr>
<th>Skill</th>
<th>Table</th>
<th>Difficulty</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Problem definition

[Dai, Mausam, Weld 2010]

Probability of answering correctly

Difficulty

High-skilled
Problem definition

[Dai, Mausam, Weld 2010]

Inference

High-skilled

Probability of answering correctly

P(True | votes) = 0.7

Difficulty
Problem definition

$T$ rounds (horizon)
Problem definition

Goal: Minimize entropy $H(A | X_S)$

$P(\text{True} \mid \text{votes}) = 0.7$

$= 0.7$

$= 0.8$

$= 0.9$

Difficulty
Offline routing
Hardness

Question utility

\[ \log \left[ \sum_{w \in \text{workers}} \gamma_w + 1 \right] \]

Sum of worker skills

\[ T = 1 \]

Skill \( \gamma \)

1

1

1

2

2
Hardness

\[ T = 1 \]

Skill \( \gamma \)

\[ \log \left[ \sum_{w \in \text{workers}} \gamma_w + 1 \right] \]

Question utility

Sum of worker skills

Theorem

NP-hard, by reduction from the PartitionProblem
The Juggler \textsubscript{OFF} algorithm

\( \mathcal{X}_S \leftarrow \emptyset \)  
while not \_finished:

\( \mathcal{X}_S \leftarrow \mathcal{X}_S \cup \{ \arg \max_{X \in \mathcal{X} \setminus \mathcal{X}_S} H(\mathcal{A} \mid x) - H(\mathcal{A} \mid X, \mathcal{X}_S, x) \} \)

Remaining assignments

Observed votes

Goal: Minimize entropy \( H(\mathcal{A} \mid \mathcal{X}_S) \)

Juggler \textsubscript{OFF}  
Submodular for known skill and difficulty
Juggler $\text{OFF}$

$T = 3$

Skill

![Diagram showing the skill level and time steps](image-url)
Juggler \text{OFF}

Reverse skill!

\[ T = 3 \]
Juggler_{OFF}

Reverse skill!

$T = 3$

<table>
<thead>
<tr>
<th></th>
<th>T</th>
<th>T</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>
Adaptive routing

Skill, Difficulty

- Latent
- Known

Policy
- Offline
- Adaptive

Concurrency
- Sequential
- Parallel
The Juggler$_{AD}$ algorithm

for $t = 1$ to $T$:

$\mathcal{X}_{S_t} \leftarrow$ call Juggler$_{OFF}$ with horizon 1

Observe votes and update prior
Juggler\textsubscript{AD}

\[ T = 3 \]

\begin{align*}
\text{T} & & \text{T} & & \text{F} \\
\text{T} & & \text{F} & & \text{F} \\
\text{F} & & \text{F} & & \text{T} \\
\end{align*}
Live experiment

12 workers

198 questions

tollner’s dismissal was announced after his team lost to Notre Dame and U.C.L.A., but he was allowed to coach through the bowl game.

Which of the following Wikipedia articles defines "Notre Dame" in exactly the way it is used in the above sentence? Tell me more

- Notre Dame Fighting Irish football

  The Notre Dame Fighting Irish football team is the intercollegiate football team of the University of Notre Dame. The team is currently coached by Brian Kelly and play home games at the campus' Notre Dame Stadium, with a capacity of 80,795. Notre Dame competes as an Independent at the NCAA Football Bowl Subdivision level and is a founding member of the Bowl Championship Series coalition (BCS). The Fighting Irish hold the highest winning percentage among college football programs and have 13 national championships recognized by the NCAA, tied for first out of all FBS schools in the post-1900 era...

- University of Notre Dame

  The University of Notre Dame du Lac (or simply Notre Dame NOH-tr-DAYM) is a Catholic research university located near South Bend, Indiana, in the United States. "Notre Dame" means "Our Lady" in French and refers to the university's patron saint, the Virgin Mary....
Live experiment
Live experiment

![Graph showing prediction accuracy vs. number of votes]

- Prediction accuracy increases as the number of votes increases.
- The graph shows a trend where accuracy improves significantly as more votes are cast.
- The label "Random" indicates a baseline or expected performance level.

Key points:
- X-axis: Number of votes
- Y-axis: Prediction accuracy
Live experiment

![Graph showing prediction accuracy vs. number of votes]

- **Round-robin**
- **Random**
Live experiment

Prediction accuracy vs. Number of votes

- Juggler_{AD}
- Round-robin
- Random
Live experiment

<table>
<thead>
<tr>
<th>Prediction accuracy</th>
<th>Number of votes</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.95</td>
<td>1500</td>
</tr>
<tr>
<td>0.9</td>
<td>1000</td>
</tr>
<tr>
<td>0.85</td>
<td>500</td>
</tr>
<tr>
<td>0.8</td>
<td>100</td>
</tr>
<tr>
<td>0.75</td>
<td>50</td>
</tr>
<tr>
<td>0.7</td>
<td>10</td>
</tr>
<tr>
<td>0.65</td>
<td>5</td>
</tr>
<tr>
<td>0.6</td>
<td>2</td>
</tr>
<tr>
<td>0.55</td>
<td>1</td>
</tr>
</tbody>
</table>

- **Juggler**
- **Round-robin**
- **Random**

52% reduction in labor for 95% of reachable accuracy
Scalability

\[ O(|\text{questions}|) \]
Binning questions

(a) Low-skilled worker

(b) High-skilled worker

High utility

Current belief

Utility

Information gain
Binning questions

Lower skill

Higher skill

Information gain

Difficult current belief

Low-skilled worker

High-skilled worker

Current belief

Lower skill

Higher skill

Information gain

Difficult current belief
Binning questions

Lower skill

Higher skill

Information gain

Difficult, uncertain

Easy, certain

Not same ordering!
Larger-scale experiment

Prediction accuracy vs. Number of votes

100 workers
2k questions
Larger-scale experiment

100 workers
2k questions

Number of votes

Prediction accuracy

20 bins
Larger-scale experiment

100 workers
2k questions

Prediction accuracy vs. Number of votes

40 bins
Larger-scale experiment

100 workers
2k questions

Prediction accuracy

Number of votes

80 bins
Larger-scale experiment

Prediction accuracy vs Number of votes

- 100 workers
- 2k questions

160 bins
Variable response times
Variable response times
Variable response times

- Juggler_{AD}
- Round-robin
- Random

Number of votes vs. Prediction accuracy
Summary

- Workers not willing to wait
- Offline algorithm Juggler\textsubscript{OFF}
- Adaptive algorithm Juggler\textsubscript{AD} Reduce labor by 52%
  - Scalability
  - Variable response times
Future directions

Skill, Difficulty

- Latent
- Known

Policy
- Offline
- Adaptive

Concurrency
- Sequential
- Parallel
Thanks!